

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/501,629  
Source: PCT  
Date Processed by STIC: 3/3/06

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 03/03/2006

PATENT APPLICATION: US/10/501,629

TIME: 12:46:28

Input Set : A:\EVRO0006.txt

Output Set: N:\CRF4\03032006\J501629.raw

```

4 <110> APPLICANT: Gurskaya, Nadejda
5      Fradkov, Arkadiy
6      Lukyanov, Sergey
7      Punkova, Natalia
9 <120> TITLE OF INVENTION: Fluorescent Protein From Aequorea Coerulscens And Uses
Thereof
12 <130> FILE REFERENCE: EVRO-0006
14 <140> CURRENT APPLICATION NUMBER: US 10/501,629
C--> 16 <141> CURRENT FILING DATE: 2004-07-15
19 <160> NUMBER OF SEQ ID NOS: 25
21 <170> SOFTWARE: FastSEQ for windows Version 4.0
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 941
25 <212> TYPE: DNA
26 <213> ORGANISM: Aequoria coerulescens
28 <400> SEQUENCE: 1
29 attcaaaaca ctgcagaatt ttggatagat tttcctgcta cttcacacgc ataaaagaca 60
30 agaaagatga gtaaaggagc agaacttttc actggagttg tccaattctc tattgaatta 120
31 aatggtgatg ttaatgggca caaattctct gtcagtggag agggcgaaagg tgatgcgaca 180
32 tacggaaagt taacctttaa atttatttgc actacaggaa aactacctgt tccatggcca 240
33 acacttgtca ctactttctc ttatggtggt caatgctttt caagatatcc agatcatatg 300
34 aaacagcatg acttcttcaa gagtgccatg cctgaagggt atatacagga aagaactata 360
35 tttttcaaaag atgacgggaa ctacaagtcg cgtgctgaag tcaagttcga aggtgatacc 420
36 ctggttaata gaattgagtt aacagggtact gattttaaag aagatggaaa catccttgga 480
37 aataaaatgg aatacaacta taacgcacat aatgtataca tcatgacaga caaagcaaaa 540
38 aatggaatca aagtttaact caaaattaga cacaacattg aagatggaag cgttcaactt 600
39 gcagaccatt atcaacaaaa tactccaatt gcgatggcc ctgtcctttt accagataac 660
40 cattacctgt ccacacaatc taccctttcc aaagatccca acgaaaagag agatcacatg 720
41 atctattttg agtttgtaac agctgctgctg attacacatg gcatggatga attatacaaa 780
42 taaatgtata gacttcaagt tgacactaac gtgtccgaac aattactaaa atctcagggt 840
43 tcctgggttaa aatcaggctg agatattatt tacatattat agattcatta gaattattta 900
44 aatactttat agatgttatt gataggggtt attttcttat t 941
46 <210> SEQ ID NO: 2
47 <211> LENGTH: 238
48 <212> TYPE: PRT
49 <213> ORGANISM: Aequoria coerulescens
51 <400> SEQUENCE: 2
52 Met Ser Lys Gly Ala Glu Leu Phe Thr Gly Val Val Pro Ile Leu Ile
53 1 5 10 15
54 Glu Leu Asn Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
55 20 25 30
56 Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
57 35 40 45
58 Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/501,629

DATE: 03/03/2006

TIME: 12:46:28

Input Set : A:\EVRO0006.txt

Output Set: N:\CRF4\03032006\J501629.raw

```

59      50      55      60
60 Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
61 65      70      75      80
62 His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Ile Gln Glu Arg
63      85      90      95
64 Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Ser Arg Ala Glu Val
65      100     105     110
66 Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Thr Gly Thr
67      115     120     125
68 Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly Asn Lys Met Glu Tyr Asn
69      130     135     140
70 Tyr Asn Ala His Asn Val Tyr Ile Met Thr Asp Lys Ala Lys Asn Gly
71 145     150     155     160
72 Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
73      165     170     175
74 Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
75      180     185     190
76 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Thr Leu Ser
77      195     200     205
78 Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Tyr Phe Glu Phe Val
79      210     215     220
80 Thr Ala Ala Ala Ile Thr His Gly Met Asp Glu Leu Tyr Lys
81 225     230     235
84 <210> SEQ ID NO: 3
85 <211> LENGTH: 717
86 <212> TYPE: DNA
87 <213> ORGANISM: Aequoria coerulescens
88 <400> SEQUENCE: 3
90 atgagtaaaag gagcagaact tttcactgga gctgtcccaa ttcttattga attaaatggt 60
91 gatgttaatg ggcacaaatt ctctgtcagt ggagagggcg aaggtgatgc gacatacggg 120
92 aagttaaccc ttaaatttat ttgcactaca ggaaaactac ctgttccatg gccaacactt 180
93 gtcactactt tctcttatgg tgttcaatgc ttttcaagat atccagatca tatgaaacag 240
94 catgacttct tcaagagtgc catgcctgaa gggtatatac aggaaagaac tatatttttc 300
95 aaagatgacg ggaactacaa gtcgcgtgct gaagtcaagt tcgaagggtga taccctgggt 360
96 aatagaattg agttaacagg tactgatttt aaagaagatg gaaacatcct tggaaataaa 420
97 atggaatata actataacgc acataatgta tacatcatga cagacaaaagc aaaaaatgga 480
98 atcaaagtta acttcaaaat tagacacaac attgaagatg gaagcgttca acttgcagac 540
99 cattatcaac aaaatactcc aattggcgat ggccctgtcc ttttaccaga taaccattac 600
100 ctgtccacac aatctaccct ttccaaagat cccaacgaaa agagagatca catgatctat 660
101 tttgggtttg taacagctgc tgcgattaca catggcatgg atgaattata caaataa 717
103 <210> SEQ ID NO: 4
104 <211> LENGTH: 238
105 <212> TYPE: PRT
106 <213> ORGANISM: Aequoria coerulescens
107 <400> SEQUENCE: 4
109 Met Ser Lys Gly Ala Glu Leu Phe Thr Gly Ala Val Pro Ile Leu Ile
110 1 5 10 15
111 Glu Leu Asn Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
112 20 25 30

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/501,629

DATE: 03/03/2006

TIME: 12:46:28

Input Set : A:\EVRO0006.txt

Output Set: N:\CRF4\03032006\J501629.raw

```

113 Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
114          35                      40                      45
115 Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
116          50                      55                      60
117 Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
118 65          70                      75                      80
119 His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Ile Gln Glu Arg
120          85                      90                      95
121 Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Ser Arg Ala Glu Val
122          100                     105                     110
123 Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Thr Gly Thr
124          115                     120                     125
125 Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly Asn Lys Met Glu Tyr Asn
126          130                     135                     140
127 Tyr Asn Ala His Asn Val Tyr Ile Met Thr Asp Lys Ala Lys Asn Gly
128 145          150                     155                     160
129 Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
130          165                     170                     175
131 Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
132          180                     185                     190
133 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Thr Leu Ser
134          195                     200                     205
135 Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Tyr Phe Gly Phe Val
136          210                     215                     220
137 Thr Ala Ala Ala Ile Thr His Gly Met Asp Glu Leu Tyr Lys
138 225          230                     235

```

141 &lt;210&gt; SEQ ID NO: 5

142 &lt;211&gt; LENGTH: 717

143 &lt;212&gt; TYPE: DNA

144 &lt;213&gt; ORGANISM: Aequoria coerulescens

146 &lt;400&gt; SEQUENCE: 5

```

147 atgagtaaag gaggcagaact ttctactgga gctgtcccaa ttcttattga attagatggt 60
148 gatgttaaatg ggcacaaatt ctctgtcagt ggagagggcg aaggtgatgc gacatacggg 120
149 aagttaaccc ttaaatttat ttgcactaca ggaaaactac ctgttccatg gccaacactt 180
150 gtcactactt tctcttatgg tgttcaatgc ttttcaagat atccagatca tatgaaacag 240
151 catgacttct tcaagagtgc catgcctgaa gggtatatac aggaaagaac tatatttttc 300
152 aaagatgacg ggaactacaa gtcgcgtgct gaagtcaagt tcgaagggtga taccctgggt 360
153 aatagaattg agttaacagg tactgatttt aaagaagatg gaaacatcct tggaaataaa 420
154 atggaataca actataacgc acataatgta tacatcatga cagacaaagc aaaaaatgga 480
155 atcaaagtta acttcaaaat tagacacaac attgaagatg gaagcgttca acttgcagac 540
156 cattatcaac aaaatactcc aattggcgat ggccctgtcc ttttaccaga taaccattac 600
157 ctgtccacac aatctaccct ttccaaagat cccaacgaaa agagagatca catgatctat 660
158 tttgggtttg taacagctgc tgcgattaca catggcatgg atgaattata caaataa 717

```

160 &lt;210&gt; SEQ ID NO: 6

161 &lt;211&gt; LENGTH: 238

162 &lt;212&gt; TYPE: PRT

163 &lt;213&gt; ORGANISM: Aequoria coerulescens

165 &lt;400&gt; SEQUENCE: 6

166 Met Ser Lys Gly Ala Glu Leu Phe Thr Gly Ala Val Pro Ile Leu Ile

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/501,629

DATE: 03/03/2006

TIME: 12:46:28

Input Set : A:\EVRO0006.txt

Output Set: N:\CRF4\03032006\J501629.raw

```

167 1           5           10           15
168 Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
169           20           25           30
170 Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
171           35           40           45
172 Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
173           50           55           60
174 Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
175 65           70           75           80
176 His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Ile Gln Glu Arg
177           85           90           95
178 Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Ser Arg Ala Glu Val
179           100          105          110
180 Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Thr Gly Thr
181           115          120          125
182 Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly Asn Lys Met Glu Tyr Asn
183           130          135          140
184 Tyr Asn Ala His Asn Val Tyr Ile Met Thr Asp Lys Ala Lys Asn Gly
185 145          150          155          160
186 Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
187           165          170          175
188 Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
189           180          185          190
190 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Thr Leu Ser
191           195          200          205
192 Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Tyr Phe Gly Phe Val
193           210          215          220
194 Thr Ala Ala Ala Ile Thr His Gly Met Asp Glu Leu Tyr Lys
195 225          230          235

```

198 &lt;210&gt; SEQ ID NO: 7

199 &lt;211&gt; LENGTH: 720

200 &lt;212&gt; TYPE: DNA

201 &lt;213&gt; ORGANISM: Aequoria coerulescens

203 &lt;400&gt; SEQUENCE: 7

```

204 atgagtaaag gaggagaact tttcactgga attgtcccaa ttcttattga attaaatggt 60
205 gatgttaatg ggcacaaatt ctctgtcagt ggagagggcg aaggtgatgc gacatacgga 120
206 aagttaaccc ttaaatttat ttgcactaca ggaaaactat gacctgttcc atggccaaca 180
207 cttgtcacta ctttctctta tgggtgttcaa tgcttttcaa gatatccaga tcatatgaaa 240
208 cagcatgact tcttcaagag tgccatgcct gaaggttata tacaggaaag aactatattt 300
209 ttcgaagatg acgggaacta caagtcgcgt gctgaagtca agttcgaagg tgataccctg 360
210 gttaatagaa ttgagttaac aggtactgat tttaaagaag atggaaacat ccttggaat 420
211 aaaatggaat acaactataa cgcacataat gtatacatca tgacagacaa agcaaaaaaat 480
212 ggaatcaaag ttaacttcaa aattagacac aacattgaag atggaagcgt tcaacttgca 540
213 gaccattatc aacaaaatac tccaattggc gatggcctg tccttttacc agataaccat 600
214 tacctgtcca cacaatctac cctttccaaa gatcccaacg aaaagagaga tcacatgatc 660
215 tattttgggt ttgtaacagc tgctgcgatt acacatggca tggatgaatt atacaaataa 720

```

218 &lt;210&gt; SEQ ID NO: 8

219 &lt;211&gt; LENGTH: 238

220 &lt;212&gt; TYPE: PRT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/501,629

DATE: 03/03/2006

TIME: 12:46:28

Input Set : A:\EVRO0006.txt

Output Set: N:\CRF4\03032006\J501629.raw

```

221 <213> ORGANISM: Aequoria coerulescens
223 <400> SEQUENCE: 8
224 Met Ser Lys Gly Ala Glu Leu Phe Thr Gly Ile Val Pro Ile Leu Ile
225 1 5 10 15
226 Glu Leu Asn Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
227 20 25 30
228 Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
229 35 40 45
230 Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
231 50 55 60
232 Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
233 65 70 75 80
234 His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Ile Gln Glu Arg
235 85 90 95
236 Thr Ile Phe Phe Glu Asp Asp Gly Asn Tyr Lys Ser Arg Ala Glu Val
237 100 105 110
238 Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Thr Gly Thr
239 115 120 125
240 Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly Asn Lys Met Glu Tyr Asn
241 130 135 140
242 Tyr Asn Ala His Asn Val Tyr Ile Met Thr Asp Lys Ala Lys Asn Gly
243 145 150 155 160
244 Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
245 165 170 175
246 Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
247 180 185 190
248 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Thr Leu Ser
249 195 200 205
250 Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Tyr Phe Gly Phe Val
251 210 215 220
252 Thr Ala Ala Ala Ile Thr His Gly Met Asp Glu Leu Tyr Lys
253 225 230 235
256 <210> SEQ ID NO: 9
257 <211> LENGTH: 717
258 <212> TYPE: DNA
259 <213> ORGANISM: Aequoria coerulescens
261 <400> SEQUENCE: 9
262 atgagtaaag gagcagaact tttcactgga attgtcccaa ttcttattga attaaatggt 60
263 gatgttaatg ggcacaaatt ctctgtcagt ggagagggcg aaggatgatgc gacatacggg 120
264 aagttaaccc ttaaatttat ttgcactaca ggaaaactac ctgttccatg gccaacactt 180
265 gtcactactc tctcttatgg tgttcaatgc ttttcaagat atccagatca tatgaaacag 240
266 catgacttct tcaagagtgc catgcctgaa gggtatatac aggaaagaac tatatttttc 300
267 gaagatgacg ggaactacaa gtcgcgtgct gaagtcaagt tcgaagggtga taccctgggt 360
268 aatagaattg agttaacagg tactgatatt aaagaagatg gaaacatcct tggaaataaa 420
269 atggaatata actataacgc acataatgta tacatcatga cagacaaagc aaaaaatgga 480
270 atcaaagtta acttcaaaat tagacacaac attgaagatg gaagcgttca acttgcagac 540
271 cattatcaac aaaatactcc aattggcgat ggccctgtcc ttttaccaga taaccattac 600
272 ctgtccacac aatctaccct ttccaaagat cccaacgaaa agagagatca catgatctat 660
273 tttgggtttg taacagctgc tgcgattaca catggcatgg atgaattata caaataa 717

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/501,629

DATE: 03/03/2006

TIME: 12:46:29

Input Set : A:\EVRO0006.txt

Output Set: N:\CRF4\03032006\J501629.raw

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date